Final Memorandum to:
Massachusetts PAs
EEAC Consultants
Copied to:
Chad Telarico, DNV GL; Sue Haselhorst ERS

From: Christopher Dyson
Date: July 17, 2018
Prep. By: Miriam Goldberg, Mike Witt, Christopher Dyson

PROJECT 73 TRACK D: EXPECTED USEFUL LIFE (EUL) ESTIMATION FOR AIR-CONDITIONING EQUIPMENT FROM CURRENT AGE DISTRIBUTION, RESULTS TO DATE

1 Introduction

One of the key research objectives of the Project 73: Track D Measure Life Method study (P73 Track D) was to explore the viability of determining measure life for some equipment using existing market research data. The P73 Track D research plan identified Project 41: The Massachusetts C&I Market Characterization On-Site Assessments and Market Share and Sales Trends Study (P41) as a potentially valuable source of information for estimating equipment measure lives. Through Project 55: The Upstream HVAC Process Evaluation (P55), DNV GL developed manufacturer nameplate research methods that further improved the usefulness of the P41 information. These P55 methods nearly doubled the percentage of HVAC units for which were able to derive equipment ages.

As part of the P73 Track D study, the DNV GL team devised a novel method for estimating air-conditioning (AC) expected useful lives (EULs) using the observed AC age data collected through the recent Massachusetts market characterization efforts (P41, P55). The team presented preliminary results in February 2018. After the Massachusetts Program Administrators (PAs) and Energy Efficiency Advisory Council (EEAC) Consultants suggested additional analysis, primarily around AC units with missing manufacture dates, the team presented revised results in April 2018.

This memorandum presents all the results to date for P73 Track D. In it, we provide a high-level description of our approach, followed by a discussion of the key data items we used from the data collection efforts, and the statistical methods we used to estimate the EULs by AC type. We present results based on the current analysis and conclude with a discussion of potential next steps.

2 Approach

This section describes our analytical approach including the data sources, key assumptions, estimation methods, and how we incorporated data into the analysis.

2.1 Data

The DNV GL team used the following data sources in the analysis:

- Observed site conditions collected in the P41 and P55 studies, including AC type and manufactured date (mostly derived from nameplate data)

- National number of AC units installed annually in the United States, from a combination of Air-Conditioning, Heating, and Refrigeration Institute (AHRI) data and U.S Department of Energy (DOE) data

2.2 Assumptions

In the analysis, we made the following key assumptions:

- That the age of an AC unit at failure/removal follows a Weibull distribution, with parameters that are constant over installation years, but may vary by AC unit type. That is, the *survival time* distribution has a Weibull structure. The Weibull is a general family of distributions, defined by a shape parameter and a scale parameter. The family includes exponential decay or constant percent "dying" in each year, corresponding to a shape parameter of 1.
- That the installation year was one year after the manufacture year of each AC unit
- That for each type of AC, the number of units installed in Massachusetts each year has been proportional to the total number of AC units installed nationally. In other words, we assumed the national data from AHRI/DOE on AC installation rates by year gives us the relative numbers installed in Massachusetts, for each AC type.

2.3 Estimation

This section describes our methods for calculating the EULs.

2.3.1 General approach

Using the assumptions described in the previous section, for any set of Weibull parameters we can determine the expected age distribution of units in place at the time of our data collection "snapshot." We start with the number of units originally installed each year, based on historical data. For any assumed survival time distribution, the number of units originally installed *y* years prior to the field work date (based on the historical installation data), combined with the probability of surviving for *y* years (from the assumed survival time distribution), tells us the number of units from the installation year that we would expect to find in place at the time of the field period. Since our analysis uses the proportions of observed units that fall into each installation year bin, rather than absolute counts, we don't need actual installation volumes; we only need the relative number installed each year.

The estimation approach is to find the Weibull parameters that minimize the least-squares difference between the observed proportions in each age bin and the expected proportions, based on the Weibull survival time distribution and the relative installation rates. Once the best-fit Weibull is determined, we take the median of that distribution as the EUL; that is, the expected age by which half of the units will have been removed.

2.3.2 Confidence intervals

Finding the best Weibull model parameters that fit the estimated age group distributions from the weighted P41/P55 effort for any AC type is a type of non-linear optimization. The translation from the Weibull parameters to the EUL is also a non-linear function. As a result, confidence interval estimation requires some form of approximation. We calculated 90 percent confidence intervals for the median using two different methods. The first is a T-test. We determine an approximate standard error of the estimated median, and calculate the confidence bounds in the usual way, as:

estimated value $\pm$ *t x standard error*

The second method is based on an F-test. We determine a 90 percent confidence region for the Weibull scale and shape parameters, based on an F-statistic, and found the extreme values of the median within that confidence region. With large enough sample sizes, these would be expected to provide similar results. In smaller samples, the T-test confidence interval can produce negative lower bounds. The F-test intervals are often asymmetric.

2.3.3 Weighting

To construct the estimated population proportion of units in each age bin, we used the installation year data from the market characterization field data, together with the sample case weights that were specifically constructed for estimation of any parameter that uses the HVAC data items collected in that study. This sample weight accounts for site-level nonresponse as well as within site, partial nonresponse (sites where only some of the HVAC data were collected). Therefore, the weighted proportions by age bin used in this analysis are full (estimated) population proportions.

Once these proportions were determined, we used two different approaches to finding the best-fit Weibull parameters. These are the parameters that, together with the assumed relative installation rates, produce expected proportions by age bin closest to the sample-based population proportions.

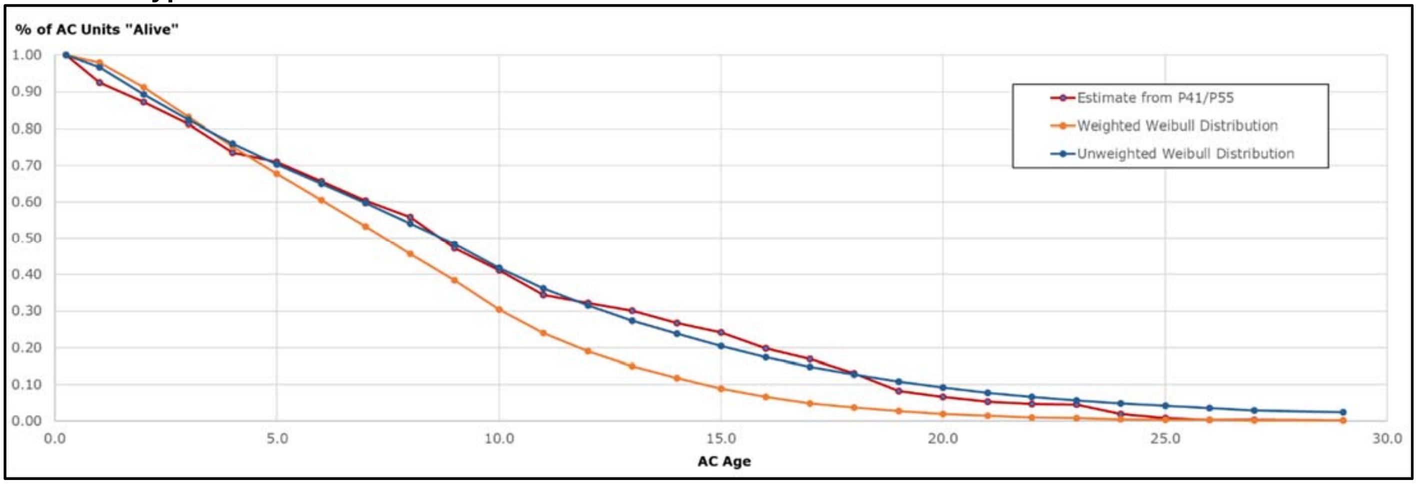
The first way to fit the Weibull accounts for the survey design-based variance and covariance of the estimated age group proportions from the market characterization study. In principle, this approach is optimal, in that it produces estimated Weibull parameters with the best precision. The second approach ignores the survey design-based variance of the age group proportions. For brevity, we refer to the first method as the "weighted" approach in the discussion below, since the age group proportions used in the Weibull model parameter estimation process were essentially "weighted" by the inverse of the survey design-based variance and covariance of the age group estimates. We refer to the second method as the "unweighted" approach, since the variance/covariance was not used. (This should not be confused, however, with the sample case weights used in the analysis. We used these sample case weights to construct weighted age group proportions for both the "weighted" and "unweighted" analyses.)

Our early analysis found that the unweighted fits tended to align better visually with the observed distributions. This is expected, because the unweighted model is assuming the age group population proportions estimated from the market characterization study are known without error. With the weighted approach, 1) age bins with a larger number of expected units have a higher variance and are therefore down-weighted in the least squares estimation; and 2) age bins with very few expected units have smaller variance and are weighted more heavily. Since typically in a survey sample, the variance (or standard error) of estimates around 50% is the greatest, these are down-weighted when finding the estimated Weibull. Since the variance (or standard error) of estimates around 5%-10% is comparably smaller, these contribute more to the estimated Weibull process. Due to these phenomena, the weighted Weibull tends to fit the age group proportions very closely at the extremes, and not as well in the middle years that are of greatest interest for this analysis.

We modified the analysis to take wider age bins at either end, so that each bin had roughly similar expected proportions, and therefore roughly similar weight. With this approach, the design-weighted and unweighted results were closer. We present the design-weighted results as preferred in most cases. However, in cases where the two results diverge more, the unweighted may still be preferred.

Figure 1 illustrates the difference between the best-fit weighted Weibull distribution and the best-fit unweighted Weibull distribution when trying to find the "best fit" for all AC units by age group. The figure shows that while the unweighted Weibull visually fits the P41/P55 better, particularly in the middle of the distribution, the weighted results fit well on the tails.

Figure 1. Comparing the Best-Fit Weighted and Unweighted Weibull Distributions for all AC Types



Note: The "basic imputation" data was used to create this figure. This is discussed in Section 4.

2.3.4 Incorporating key data sources

This subsection describes our approaches for incorporating the key data sources into the analysis.

2.3.4.1 Install date versus manufacture date

The data we had for each unit was usually the manufacture date, determined from the nameplate data. We measured the observed "age" as the time between the installation date and the site observation date, assuming a one-year lag from manufacture year to installation year.

2.3.4.2 Installation rate

AHRI provided data on the number of installations of unitary AC units from 1997 through 2016.[1] DOE[2] showed installation rates of split and packaged AC from 1953 to 2010. We fit a regression to predict AHRI as a function of the sum of split and packaged units from DOE, using the overlapping years. We used that

[1] AHRI, historic AHRI shipment data are from http://www.ahrinet.org/Resources/Statistics/Historical-Data.aspx, and monthly data since 2010 are at http://www.ahrinet.org/statistics.aspx.

[2] US Department of Energy Office of Energy Efficiency and Renewable Energy, August 2015. TECHNICAL SUPPORT DOCUMENT: ENERGY EFFICIENCY PROGRAM FOR CONSUMER PRODUCTS: Residential Central Air Conditioners and Heat Pumps.

relationship to project the AHRI series to early years. This approach applies the single installation rate pattern we developed to all regions and all types of units.

2.3.4.3 Missing manufacture year

After the DNV GL team had completed look-ups for all sites with legible nameplate pictures, about one third of the records in the market characterization data set still had missing manufacture years. For the records with a missing year, we imputed the installation year using the respondent data, including any information a respondent might have provided, such as a range for estimated installation year.[3] The imputation accounted for varying age by AC equipment type, business type (education, food sales, etc.), and annual kWh consumption categories (3 levels). We refer to this approach as the "basic imputation" of age.

The purpose of the imputation was to adjust for the possibility that units with a missing age tend to be older than units with an identifiable age. The basic imputation increased the age distribution only slightly, resulting in about a half year increase to the EUL.

If older units are more likely to have a missing age, the "basic" imputation doesn't fully account for that, as it assumes that the age distribution for those with missing age is the same as the distribution for those with known age, accounting for variations by building type, consumption group, and "guess" bin. On the other hand, there is no way to address the possibility that older units have a different distribution than what the age-respondent data suggests, without additional information or additional assumptions.

To illustrate a potential additional assumption, we constructed an alternative set of imputations with the added assumption that all ACs with unknown age had a manufacture year of 2008 or earlier. Thus, with a field period of 2014-2015, the units with unknown age were assumed to be at least 6 or 7 years old. This cut-off was assumed because it corresponded roughly to the initial EUL estimates from preliminary analysis prior to imputations. Results using both the basic imputation and the alternative imputation are presented in the next section.

3 Results

This section presents the results from our EUL analysis.

3.1 Age imputation

shows the average unit age for the data set with known manufacture year, the full data set including the basic imputation, and the full data set with the alternate imputation. All the averages shown are sample-weighted; that is, these are the estimated population average ages of equipment in place at the time of the field data collection. The basic imputation increases the average age by about half a year compared to excluding the cases with unknown year. The alternative imputation increases average age by nearly two years compared to excluding the unknown year cases.

[3] As part of the P41 onsite survey, the DNV GL team asked customers to estimate whether their AC equipment had been installed "sometime after 2008" or was 2008 or older. So, in many cases where the unit's manufacture date is missing, we still have the customer-estimated age. The P55 study found that when it compared these customer self-reports of equipment age ranges to the actual equipment ages based on the manufacturer nameplate information, the customer self-reports were very accurate.

Table 1. Average observed unit age before and after age imputation

AC type	Excluding cases with unknown year	Including cases with unknown year	
		Basic imputation	Alternative imputation
Split system AC condensing unit	9.1	9.3	10.4
Split system heat pump	5.7	6.8	7.4
Package RTU AC	9.0	9.5	11.1
Package system heat pump	7.4	8.0	8.9
Mini split AC	8.5	8.9	10.0
Packaged terminal AC (PTAC)	7.3	8.4	9.7
Mini split heat pump	11.5	11.5	12.7
Total ACs	8.5	9.1	10.3

Figure 2 and Figure 3 below show how the age distribution is shifted with the two imputation approaches.

Figure 2. Percent of units in each age bin, without imputation (respondents) and with basic imputation (respondents and non-respondents combined)

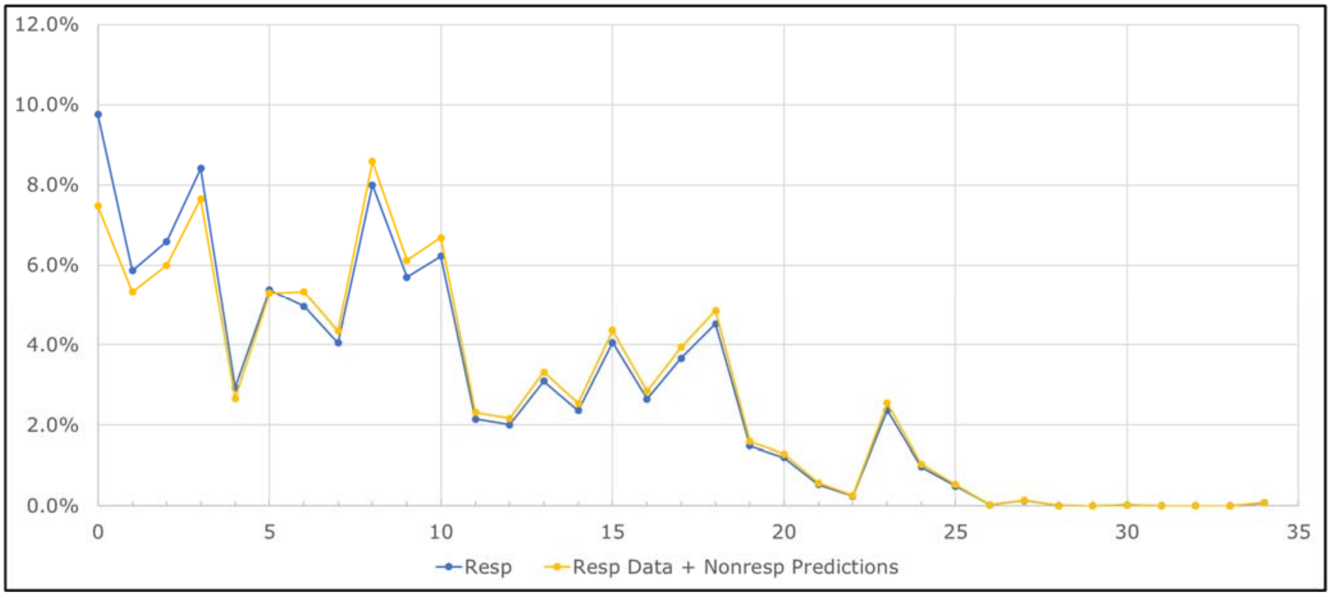


Figure 3. Percent of units in each age bin, without imputation (respondents) and with alternative imputation (respondents and non-respondents combined)

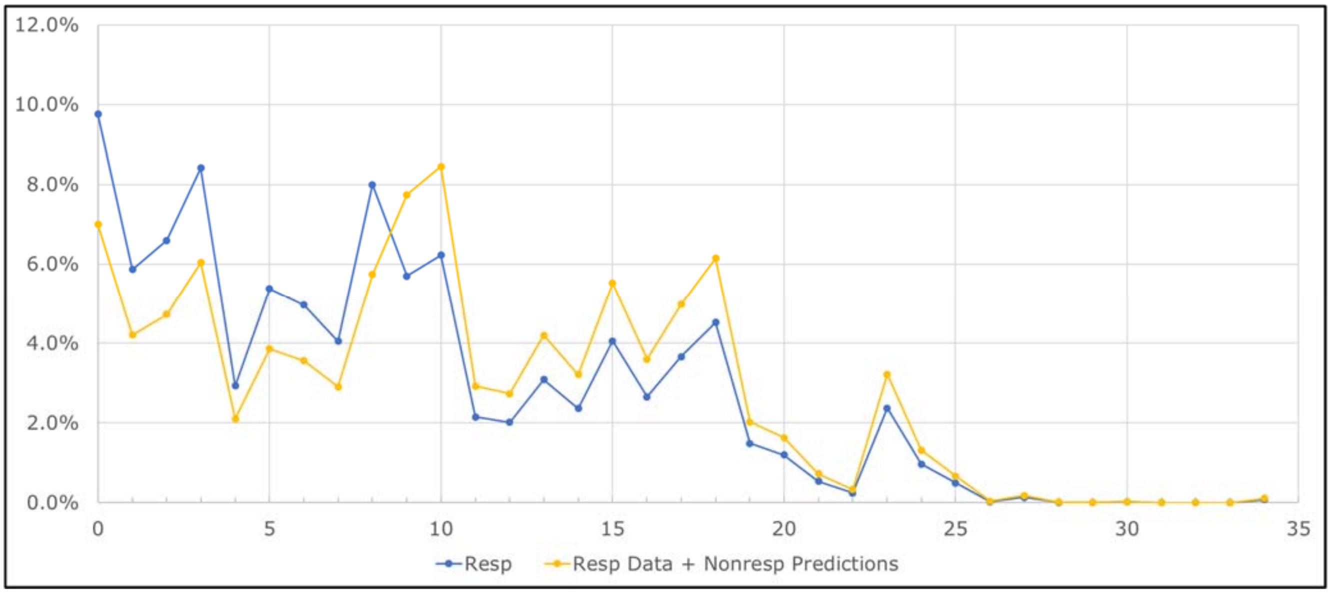


Table 1, Figure 2, and Figure 3 show the distributions of ages of equipment in place at the time of the market characterization field visits. By contrast, the estimated Weibull distributions presented next are the distributions of the ages of equipment at failure or removal.

3.2 Median lifetime estimates

Table 2 below shows the estimated EULs with the different imputation methods. The EUL is calculated as the median time to replacement or removal, based on the estimated Weibull survival time distribution. Median time in place and operational is a common definition of the EUL. The estimated medians are the median ages determined from the fitted Weibull curves. This table presents the weighted Weibull results.

Table 2. EUL estimates with different imputation methods (weighted Weibull fit)

AC type	Excluding cases with unknown year	Including cases with unknown year	
		Basic imputation	Alternative imputation
Split system AC condensing unit	7.1	7.1	7.1
Split system heat pump	n/a	n/a	n/a
Package RTU AC	8.0	8.9	9.3
Package system heat pump	6.3	6.7	8.2
Mini split AC	3.7	7.1	10.5
Packaged terminal AC (PTAC)	4.6	8.5	22.4
Mini split heat pump	37.2	35.1	19.9
Total ACs	**5.6**	**6.0**	**7.5**

Table 3 shows more complete distributional characteristics of the fitted Weibull curves for each AC type. The statistics shown are the mean and quartiles of the time to failure or removal, based on the fitted Weibull survival time distribution. Also shown are sample sizes, and confidence intervals for the median.

Several observations can be made from the table:

- The median lifetimes are mostly in the range of 7 to 9 years across the different AC types.
- Mean lifetimes are greater than medians, as is expected for age distributions. However, means can take on more extreme values.
- For AC types with fewer than 100 sites, the confidence bounds tend to be fairly wide. For AC types with 300 or more sites, the confidence bounds are tighter and more similar between the two estimation methods.
- The fit for all AC types together produces median lifetimes smaller than do any of the individual types alone. Reasons for this are unclear.

Table 3. Numbers of cases and distributional statistics of the best-fit Weibull survival time distribution, basic imputation

Technology	Fit	Total AC units in field sample	Total sites with AC unit in sample	Q1	Median (EUL)	Mean	Q3	SE(Med)	90% confidence interval for median			
									Est-t*SE	Est+t*SE	F-Test Lower	F-Test Upper
All AC types	Weighted	10,197	618	3.3	6.0	7.1	9.8	0.5	5.1	6.9	4.5	7.4
	Unweighted	10,197	618	3.4	7.4	9.8	13.6	2.5	3.2	11.6	0.0	17.5
Split system AC condensing unit	Weighted	1,702	316	5.8	7.1	7.0	8.3	0.2	6.7	7.5	6.6	7.6
	Unweighted	1,702	316	3.8	7.8	9.9	13.8	1.1	6.0	9.7	0.0	9.4
Split system heat pump	Weighted	398	57	*	*	*	*	*	*	*	*	*
	Unweighted	398	57	1.4	14.8	166.4	96.1	16.5	-14.6	44.1	0.0	25.3
Package RTU AC	Weighted	2,150	369	5.5	8.9	9.6	12.9	0.5	8.0	9.7	7.3	10.4
	Unweighted	2,150	369	4.1	8.1	9.9	13.7	1.2	6.0	10.1	2.4	15.3
Package system heat pump	Weighted	248	26	5.7	6.7	6.7	7.8	0.2	6.3	7.2	4.9	27.7
	Unweighted	248	26	5.5	6.5	6.5	7.6	0.3	5.8	7.3	0.0	32.4
Mini split AC	Weighted	702	196	2.1	7.1	14.5	18.5	2.3	3.1	11.1	3.5	21.7
	Unweighted	702	196	4.7	7.8	8.6	11.6	1.4	5.4	10.2	0.4	29.6
Packaged terminal AC (PTAC)	Weighted	4,271	53	2.3	8.5	19.5	24.0	3.1	2.7	14.4	4.0	23.7
	Unweighted	4,271	53	2.2	7.2	14.5	18.7	2.3	2.9	11.6	3.8	23.7
Mini split heat pump	Weighted	726	45	25.7	35.1	35.6	44.9	0.1	34.9	35.3	0.9	38.6
	Unweighted	726	45	6.1	23.0	53.7	65.6	46.2	-64.7	110.6	0.9	38.6

Note: *Weibull parameter estimation process did not yield reasonable results.
Q1 and Q3 are the 25th and 75th quartiles of the estimated Weibull distribution.

4 Potential next steps

4.1 Outstanding concerns

EULs on the order of 7-9 years are lower than the current 15 years in the MA TRM. The MA AC EULs prior to 2005 averaged around 9.5 years. (ERS, 2005) The 2005 ERS report recommended that the current values be around 15 years based on comparison with values used around the country.

The EEAC and the PAs provided comments on the draft version of this memorandum. They and the program evaluators also discussed these findings in two different Massachusetts C&I Evaluation Team weekly meetings in July 2018. In their comments on the memo and during these weekly meetings, some of the EEAC and PA representatives expressed scepticism about measure lives as low as 7-9 years, especially for certain types of AC equipment. For example, while they indicated that an EUL around 7-9 years might be plausible for PTACs, which tend to have large numbers at a site and may get replaced in bulk, they thought this less likely for roof-top units.

If the results found here are systematically low (in other words, they are biased and under-estimate the true EULs in the population), this could have a few different causes:

1. The age imputation doesn't adequately account for the greater tendency of older units to have unknown ages in the market characterization data. In other words, the actual age of AC units with missing age is larger than what the respondent data indicates.
2. The national installation rates may not reflect Massachusetts patterns, or those of each technology.
3. The market characterization data from the P41/P55 studies under-represents older businesses or businesses in older buildings for some reason not accounted for by the sample weights.

4.2 Potential next steps

If the PAs wish to pursue this analysis further, below are some possible next steps.

4.2.1 Address age imputation more thoroughly

The imputations could be repeated with alternative assumptions, taking into account additional information on the site, as well as the reasons for missing age data, and making some corresponding assumptions.

After the analysis described above was completed, we reviewed the reasons some units have missing age, and identified a set of assumptions that could be appropriate for each of these. Table 4 lists the prevalence of each condition, the associated assumption for imputation, and a specific assumption that could be applied in the analysis. These specifics are for illustration only and could be assigned differently.

Table 4. Reasons for missing age data and potential treatment in analysis

Reason for no age data	Approximate % of missing records	Imputation assumption	Illustrative specific assumption
Nameplate weathered	10%	Age distribution older than non-missing	> 10 years
Valid legible number, not found in data bases	15%	Some are obscure, some are old	50% same as non-missing, 50% >15 years
No information	11%	Equipment inaccessible	Same as non-missing
Partial information, no picture taken, no date found on equipment	66%	Sometimes age related but mostly not	67% same as non-missing, 33% >10 years

4.2.2 Addressing installation rates

One way to address the possibility that the national data is not representative of Massachusetts is to purchase regional data of the same type. However, after further exploration of this option, the evaluation team learned that Massachusetts-specific and regional-specific C&I HVAC sales data is only available for the 2013-2018 period and is very expensive. We could also run scenarios with alternative installation rate assumptions, perhaps adjusting national or regional installation rates based on national and Massachusetts levels of new construction.

4.2.3 Possible under-representation of older buildings in the field data sets

We have no reason to believe there was under-representation of older buildings in the P41/P55 field data, other than the finding of overall lower EULs, which reflects a younger distribution of existing units than some reviewers expected. We did not have more difficulty recruiting in older areas than in newer ones. Since age was not available for sampling purposes, there was no stratification to control for the age distribution. There was also nothing in the sampling and recruiting that would lead us to expect any bias in this respect.

To assess whether there is any systematic age bias in the sample, we would need an alternative source of population building age distributions. The customer profile data being developed has "year constructed" attached to about two-thirds of the C&I accounts. This information comes from tax assessor's data. Given this level of missing data, which is similar to the proportion of missing age in the equipment data, the available building age data may not be a good indicator of how representative our equipment age records are.

Alternatively, if we have premise in-service data from the utilities, this could be an alternative or supplemental source. If we did have a reasonably complete set of age data for the 2014-15 accounts, we could assess whether the P41/P55 samples tend to have younger buildings than the general C&I population. If there is evidence of such an effect, we could post-stratify the sample to calibrate to the known age distribution.

5 Recommended EUL values for C&I unitary HVAC equipment

In the two Massachusetts C&I Evaluation Team weekly meetings in July 2018, the EEAC and the PAs came to an agreement that the current EULs for C&I unitary HVAC equipment be reduced from the current TRM value of 15 years to 12 years based on the existing evidence. This EUL reduction will be one of the new C&I parameters in the 2019-2021 Massachusetts plan. This decision was based on the following considerations:

- EEAC/PA skepticism about whether the EUL values of 7-9 years that came out of the analysis described in this memo were realistic based on their knowledge of the Massachusetts C&I HVAC market and C&I building stock
- Some of the limitations of the analytical method described in this memorandum, such as the estimation of Massachusetts equipment installation rates based on national sales data, and the fact that about a third of the HVAC equipment from the P41 and P55 databases were missing manufacture date information
- The acknowledgement that the 15-year EUL estimate currently in the Massachusetts TRM came from a 2005 ERS study which likely derived it from an average of EULs from TRMs across the country. This 2005 study did not critically examine the sources of these national EUL estimates. The P73 Track D literature review also raised doubts about the reliability of many EUL estimates in TRMs.[4]
- The fact that the evaluators could not find many units older than 15 years in the P41 and P55 databases, with the acknowledgement that many of these older units may be represented by the HVAC equipment which lacked manufacturer dates.

During these July 2018 meetings, the EEAC and the PAs also agreed that this new 12-year EUL estimate was preliminary based on the existing evidence. They recommended that some of the additional research suggested in this memo be pursued in the coming months to improve the reliability of this estimate.

[4] "DNV GL's review of conference and white papers revealed a general lack of recent primary research on measure lives, especially involving methodologies for estimating them. Most recently published work details common failings of measure life estimates in TRMs and recommends best practices for performing measure life studies and choosing estimates. Several studies by Skumatz et al. point out that most EUL values used in energy efficiency programs are not based on original research or independent studies. Instead, they are based on existing tables from other jurisdictions (or summary studies of several tables). Many of these tables lack proper citations of source documents, and when they do cite their sources, the citations are surprisingly circular. In many cases, the TRMs do not cite their sources. Skumatz et al. also identified other specific issues with measure life tables and values, including: A lack of differentiation for common variables such as program design, climate, and other operating conditions." (Project 73 Task D: TRM/Literature Review Key Findings Memorandum, March 19, 2018, p. 2.)